Natural and Anthropogenic Factors Influencing Freshwater Fish Species Diversity

A Research Plan

Denis White
Western Ecology Division
National Health and Ecological Effects Research Laboratory
Office of Research and Development
US Environmental Protection Agency
200 SW 35th Street
Corvallis, Oregon 97330 USA

Date of plan: 20 October 1999

Notice. The research described in this plan will be funded by the US Environmental Protection Agency. This document has been subjected to the Agency's peer and administrative review and approved for publication. Mention of trade names or commercial products does not constitute endorsement nor recommendation for use.

1 Introduction

The United States has an impressive diversity of freshwater fish, about 800 species, or about 10% of the known species in the world, and is the seventh richest country in this respect (Master et al. 1998). But more than 300 of these species are at risk of extinction, and 17 species are already known to be extinct (Stein and Flack 1997). Furthermore, this accounting only covers the species level of organization; on March 16, 1999, nine additional “evolutionarily significant units” (genetically distinct populations, stocks, or runs; Waples 1995) of Pacific Northwest salmonid species were listed by the National Marine Fisheries Service as either threatened or endangered under the Endangered Species Act (http://www.nwr.noaa.gov/1salmon/99LIST2a.gif). With these new postings there are now 24 Northwest salmonid stocks listed.

People value fish for many reasons, and in the United States we have empowered the government to take major actions to help conserve these animals through the Endangered Species Act and other legislation. Although we have considerable knowledge about the numbers of fish species and their geographical distributions, we know much less about why there are such numbers of species, why they are so distributed, and how human activities may affect these distributions. In order to develop successful conservation policies for fish biota, any additional understanding we can gain about these questions can only make these policies more sound.

This research plan proposes to look at the distribution of fish species in space and time using two major approaches, (1) an empirical study of contemporary distributions on the west and east coasts of the United States looking at potential environmental correlates of these distributions, and (2) a simulation study of the dynamics of species distributions in stream networks. These studies hope to make a small contribution to answers to these fundamental questions:

Why is the diversity of fish species distributed as it is?
What are the apparent effects of human activities on these distributions?

Looking at these questions through the analysis of contemporary data has the potential to provide complementary and useful information to an approach based on simulation modeling. The empirical data analysis can help to parameterize the modeling formulation, and the modeling may suggest other variables to include in the empirical data analysis.

This plan introduces the problem and describes the two approaches in detail.

1.1 A classical problem - why just so many species?

One of the great questions about the natural world is why there are so many species, or, more comprehensively, why there are just so many species, and not a lot more nor a lot less. This question, at the heart of natural history, biogeography, and community ecology, is implicit if not explicit at least as far back as the work of Linnaeus and Buffon (Brown and Lomolino 1998). Currie (1991) stated the question in this way: “why, in a given environment, does one find *n* different species, not more or fewer, and why does *n* vary so greatly among environments?” (p 27). This rephrasing of the title question of Hutchinson’s famous Santa Rosalia address (1959) helps in developing a quantitative approach to the question.

Since Hutchinson's conjectures, investigators have proceeded in several ways. Modeling studies of resource partitioning (for example, May and MacArthur 1972, Maiorana 1978, Rappoldt and Hogeweg 1980) attempted to support Hutchinson's hypothesis that a constant ratio in the size of foraging equipment traits was necessary for sympatric species to occupy different niches at the same level in a food web hierarchy. Such a fixed constant distance would provide an upper bound on the number of species in a given environment having a known amount of resources. Further study by Heck (1976) and Roth (1981), however, were critical of, or very cautious with, these model results, and Simberloff and Boecklen (1981) showed little support for the fixed distance hypothesis from a statistical analysis of available data sets. Furthermore, a critic with an allometric point of view, Reiss (1989), argued that a weakness in the modeling studies was not accounting for the dependence of rates of reproduction on size of organism, and the consequences of this dependence on coexistence.

Brown's (1981) revisiting of Hutchinson's question suggested changing the focus to species-energy relationships: how do species divide available energy, or, alternatively, how are species numbers related to productivity? Wright (1983) introduced the concept of species-energy curves as an alternative to species-area curves to model the hypothesized relationship between number of species and amount of available energy per unit area. Large area studies by Currie and Paquin (1987; trees), Currie (1991; trees, birds, mammals, amphibians, reptiles), and Fraser and Currie (1996; corals) have supported this hypothesis. Kerr and Packer (1997; mammals) found less support.

Simpson, in a study of North American mammals (1964), stimulated a group of studies that focused on large area patterns of terrestrial and freshwater species diversity as represented on geographic grids. Two successor studies to Simpson were Cook (1969; birds) and Kiester (1971; reptiles, amphibians). These studies included speculation on possible mechanisms influencing these distributions. In succeeding years a number of investigators have studied such patterns and possible environmental correlates using quantitative methods: Wilson (1973; mammals), Rogers (1976; reptiles, amphibians), McAllister et al. (1986; fish), Owen (1989, 1990a, 1990b; mammals, reptiles, amphibians), McAllister et al. (1994; fish), Shvarts et al. (1995; mammals), O'Connor et al. (1996: birds), Böhning-Gaese (1997; birds), Rathert et al. (in press; fish), plus studies listed in the preceding paragraph.

A number of studies have looked specifically at species diversity of fish, for example: Horwitz (1978), Livingstone et al. (1982), Legendre and Legendre (1984), Eadie et al. (1986), Hughes and Gammon (1987), Hughes et al. (1987), Beecher et al. (1988), Matuszek and Beggs (1988), Angermeier and Schlosser (1989), Hugueny (1989), Jackson and Harvey (1989), Osborne and Wiley (1992), Oberdorff et al. (1993), Magnuson et al. (1994), Hugueny and Paugy (1995), Mandrak (1995), Oberdorff et al. (1995), Kelso and Minns (1996), Griffiths (1997), Oberdorff et al. (1997), Whittier et al. (1997), Guégan et al. (1998), Oberdorff et al. (1998), Angermeier and Winston (1999), Allen et al. (in press), and Whittier and Kincaid (in press), in addition to the McAllister et al. and Rathert et al. papers listed above. Species-area effects, stream order, and other historical and environmental factors were considered in these studies (Table 1).

Table 1. 28 studies of fish species diversity. The first column lists the first author, except for Hughes et al. where the two papers by Hughes from 1987 are distinguished. The second column is the date of publication. The third column indicates whether the samples were collected at sites and related to a corresponding lake, drainage area, or basin, or assigned to a regular geographic grid. The fourth column is the number of samples. The fifth column indicates some of the main relationships investigated about fish species richness or composition. For some studies, e.g., Magnuson, the brief description of relationships does not include some of the significant issues addressed.

First Author	Date	Site/ Grid	N	Relationships
Allen	in press	Site	186	Species richness vs. other taxa, lake morphology, littoral structure, climate, water quality, human development
Angermeier	1989	Site	200	Species richness vs. abundance, habitat volume, habitat area, habitat complexity, habitat type
Angermeier	1999	Site	348	Species, guild composition vs. basin identity, physiography, stream order, elevation, slope, map coordinates
Beecher	1988	Site	71	Species richness vs. elevation, slope, basin area, stream order
Eadie	1986	Site	82	Species richness vs. lake area, latitude, longitude
Griffiths	1997	Site	651	Species richness vs. lake area, depth, elevation
Guégan	1998	Site	183	Species richness vs. basin area, discharge, terrestrial productivity
Horwitz	1978	Site	1065	Species richness vs. basin position
Hughes	1987	Site	26	Species richness, composition vs. physical-chemical habitat, stream order
Hughes et al.	1987	Site	85	Species composition vs. ecoregions, basins, physiographic regions
Hugueny	1989	Site	26	Species richness vs. discharge, basin area, vegetation
Hugueny	1995	Site	47	Species richness vs. regional species pool, season, basin position
Jackson	1989	Site	286	Species richness, composition vs. lake area, depth, pH, region
Kelso	1996	Site	27	Species composition vs. contaminants, water quality
Legendre	1984	Grid	289	Species composition vs. region, climate, interbasin connectivity, glacial history, vegetation
Livingstone	1982	Site	26	Species richness vs. discharge, basin area, basin length
Magnuson	1994	Site	7	Species richness vs. lake area
Mandrak	1995	Grid	80	Species richness vs. postglacial dispersal, climate, elevation, lake chemistry, lake morphometry
Matuszek	1988	Site	2931	Species richness vs. lake area, depth, elevation, lake chemistry, latitude, longitude
McAllister	1986	Grid	907	Species richness vs. climate, glacial history, basin identity
Oberdorff	1993	Site	55	Species richness, guilds vs. basin area, basin position
Oberdorff	1995	Site	292	Species richness vs. basin area, discharge, latitude, climate, terrestrial productivity, glacial history, continental richness
Oberdorff	1997	Site	132	Species richness vs. discharge, basin area, climate, terrestrial productivity, latitude, peninsular status, basin position, glacial history, refugial distance
Oberdorff	1998	Site	9	Species richness, density, biomass vs. ocean distance, elevation, stream width, discharge, basin area, source distance, regional richness
Osborne	1992	Site	48	Species richness, composition vs. basin area, stream order, discharge, basin position
Rathert	in press	Grid	375	Species richness vs. climate, stream density, lake area, interbasin connectivity, ocean distance, road density, human density, exotic species
Whittier	1997	Site	195	Species richness vs. lake area, local human disturbance, watershed human disturbance
Whittier	in press	Site	203	Species richness vs. lake area, depth, elevation, exotic species, human disturbance

Ricklefs (1987), plus authors of chapters in Ricklefs and Schluter (1993), and Cornell (1997) looked at the relative influence of local and regional processes on species diversity in communities. These reviews and studies covered many possible causes for diversity from biotic interactions to environmental and historical factors, but with an explicit focus on composition and diversity of local communities.

One other major line of inquiry into species numbers is island biogeographic theory (MacArthur and Wilson 1967, Brown and Lomolino 1998). In addition to observational and experimental studies, this theory has fostered a number of modeling studies using a spatial framework for the environment: Seagle and Shugart (1985) on animals colonizing islands undergoing habitat transitions; Seagle (1986) refining the preceding study; Villa et al. (1992) on animal demographics influencing colonization of islands; Haydon et al. (1993) on the relative influence of ecology, geological history, and chance on colonization and speciation of islands; and Hubbell (1994) on the “community drift” model of species diversity and abundances in space. In related studies, Gore and Milner (1990) discussed colonization models based on island biogeographical theory for stream systems; Wright et al. (1993) modeled local and regional species richness based on the statistics of regional richness deriving from that of local incidence which, in turn, was derived from sufficient local abundance resulting from energy partitioning; Koopowitz et al. (1994) developed a related type of model for extinction under habitat loss; Kaufman et al. (1998) used trophic relationships over space to look at dynamics of extinction; and Ziv (1998) developed a comprehensive spatial model incorporating local ecological effects with global colonization processes for studying diversity patterns.

1.2 Two approaches - empirical associations and simulation modeling

Studying a problem as complicated as species diversity patterns offers many possible approaches. A better perspective on a problem can come both from understanding in depth from a single point of view, and from understanding in breadth from a number of points of view. This research takes the position of attempting to approach the problem from two relatively different positions while focusing on only one major taxon, fishes, for which there is some knowledge about species diversity but also many unanswered questions.

First, the intention is to investigate empirically mechanisms that could explain fish species distributions using the best available data on these distributions over large areas of the US. In particular this research plans to use distribution data prepared by The Nature Conservancy (TNC) (Master 1994, 1996) for both native and exotic fish species in the three west coast states of the US plus five east coast states in the Chesapeake Bay region. These data purport to be comprehensive in taxonomic coverage and relatively fine in spatial resolution (over a regular grid with cells of about 650 km^2 in size). This effort will build on a similar study of a more restricted geographic scope (Rathert et al. in press) and broaden both the spatial scope as well as possible explanatory mechanisms. Data for fish species response and potential predictive factors will be analyzed with several statistical techniques.

Supplementing these analyses will be an effort to compare results from these data with analyses of data from the Environmental Monitoring and Assessment Program (EMAP) on fish species distributions from the middle Atlantic highlands sampling project, and from west coast sampling projects as these data are or become available. From these comparisons, carried out in collaboration with EMAP scientists, the research effort will, first of all, look at the strength of relationships obtained from the TNC data, and, secondly, examine the relative merits of these complementary datasets.

From these studies the hope is to learn how patterns of possible forcing functions are related to patterns in fish species diversity. Understanding empirically the relative importance of climatic factors versus habitat heterogeneity versus anthropogenic disturbance, for example, can suggest possible approaches for conservation or protection activities. As correlative statistical studies, the results will be likely to confirm some existing hypotheses and contradict others. Those hypotheses best supported can be further examined by simulation modeling.

In simulating the occupancy of a stream network by assemblages of fish species, the intention is to learn more about mechanisms influencing the distribution of these assemblages. By incorporating processes such as colonization, niche fitness, interspecific competition, population growth, and local extinction, this work will attempt to mimic these factors as they are understood to affect fish assemblages. The focus of this work will not be at the level of individual organisms but rather at the level of a collection of individuals occupying a link in a network. By abstracting to this level of organization it is hoped that the interactions and collective patterns of multiple species can be simulated, a goal that is difficult with single species population dynamics simulations.

The outcomes of the simulation modeling may provide insight into what factors influencing species diversity should be investigated more closely in empirical studies. This feedback between these two aspects of this research plan offer synergistic research possibilities.

1.3 Why is this research relevant to the Environmental Protection Agency?

The US Environmental Protection Agency (EPA) bases much of its activities on laws and regulations of the US that have been enacted to protect the environment. At the level of congressional legislation, more than a dozen laws enable EPA’s activities (http://www.epa.gov/epahome/laws.htm). For protecting the fish diversity of the US, EPA activities reference three major laws. Arguably the most important of these is the Clean Water Act (CWA) some of whose goals and provisions date as far back as the 1948 Federal Water Pollution Control Act (http://www.epa.gov//epahome/lawintro.htm). The most recent wording and authorization of the CWA occurred in 1987 (http://www.epa.gov//region5/defs/html/cwa.htm). This legislation is part of the United States Code, Title 33 - Navigation and Navigable Waters, Chapter 26 - Water Pollution Prevention and Control, Sections 1251 - 1387 (legal citation is 33 USC §§ 1251 - 1387; text is at http://www4.law.cornell.edu/uscode/33/ch26.html).

The congressional declaration of goals and policies in section 1251 of the CWA says that “it is the national goal that wherever attainable, an interim goal of water quality which provides for the protection and propagation of fish, shellfish, and wildlife and provides for recreation in and on the water be achieved ...” The EPA is assigned to administer this chapter of the code.

A second law is the National Environmental Policy Act (NEPA; 42 USC §§ 4331 - 4335; text at http://es.epa.gov/oeca/ofa/nepa.html). NEPA says in section 4331 that “it is the responsibility of the federal government to use all practicable means” to have federal actions planned and coordinated such that “the Nation may (1) fulfill the responsibilities of each generation as trustee of the environment for succeeding generations ...”

EPA has an important role under NEPA to review all environmental impact statements (EISs) prepared by federal agencies. Although there is no explicit mention of fish (or other wildlife taxa) in the NEPA law or in the main regulations and guidance for implementing it (see http://ceq.eh.doe.gov/nepa/nepanet.htm), the goal quoted above is surely so comprehensive as to include fish and wildlife as a component of the environment for which the federal government should assist in the maintenance.

The third law is the Endangered Species Act (ESA; 16 USC §§ 1531 - 1544; text at http://www4.law.cornell.edu/uscode/16/ch35.html). Section 1531 states that “the Congress finds and declares that (1) various species of fish, wildlife, and plants in the United States have been rendered extinct as a consequence of economic growth and development untempered by adequate concern and conservation; (2) other species of fish, wildlife, and plants have been so depleted in numbers that they are in danger of or threatened with extinction; (3) these species of fish, wildlife, and plants are of esthetic, ecological, educational, historical, recreational, and scientific value to the Nation and its people.” The policy enabled in this section is that “all Federal departments and agencies shall seek to conserve endangered species and threatened species ...” An important part of the law is in section 1532 under definitions that says “(16) The term ‘species’ includes any subspecies of fish or wildlife or plants, and any distinct population segment of any species of vertebrate fish or wildlife which interbreeds when mature.” It is under this definition that populations or stocks of salmonids have been listed as described in the introduction to this document.

EPA has the responsibility under ESA to consult with the managing agency (US Fish and Wildlife Service) on issues concerning states’ implementations of the Clean Water Act and the pesticide registration process as these activities affect the implementation of ESA. In addition, the EPA administrator is one of seven members of the committee (the so called “God Squad”) that considers exemptions from the ESA. See http://www.fws.gov/r9endspp/section7/section7.htm for documents pertaining to this consultation responsibility.

While these major environmental laws of the United States involve the EPA in somewhat different ways, it is clear that the government as a whole is committed to the protection and maintenance of fish species of this country, and that EPA, especially through administration of the Clean Water Act, has a major responsibility. Angermeier and Schlosser (1995), and references cited therein, have argued that, in fact, the CWA has the important advantage of covering all aquatic systems, not just habitats for endangered or threatened species as the ESA does.

Not just legislation requires or urges EPA to protect fish diversity. In an important document issued by EPA’s extramural public advisory group, the Science Advisory Board (EPA - SAB 1990), “habitat alteration and destruction” and “species extinction and overall loss of biological diversity” were listed as the first two of the four relatively high-risk problems that the SAB considered most important for EPA to address in its research and protection activities. This statement by the SAB helps to construct an inclusive focus for protection of fish species diversity that incorporates both the specific and general concerns of legislation and the relative risk rankings of the SAB.

In order to protect fish (and wildlife) diversity, it is necessary to address more than the habitat for a single species. It is necessary to understand as much as possible about the behavior and distributions of the collected fish fauna of the country over river and stream networks large and small. Although studies in specific places and on specific taxa contribute essential information to this understanding, it is important that some effort be allocated to looking at the empirical distributional facts of the fish species of the country over large areas and to attempting to understand the dynamics of how these fish so distribute themselves. This is an effort that this research plan proposes.

1.4 How this research fits into team, branch, and division objectives

The proposed research will be conducted at the Western Ecology Division of EPA's National Health and Environmental Effects Laboratory, as part of the Regional Ecology Branch's Landscape Team. The work will be performed in Corvallis, Oregon.

The Western Ecology Division has the mission "(1) to provide EPA with national scientific leadership for terrestrial and regional-scale ecology, and (2) to develop the scientific basis for assessing the condition and response of ecological resources of the western United States and the Pacific Coast"; the division research approach comprises: "(1) developing an understanding of the structure and function of ecological systems, and (2) conducting holistic analyses of ecological phenomena at the ecosystem, landscape, and regional scales" (http://www.epa.gov/wed/pages/mission.htm). The research proposed here is innovative in design and focused on understanding an important part of aquatic system response. The empirical work explicitly focuses on the west coast, as well as part of the east coast for comparative purposes; the simulation modeling hopes to foster an approach that can be adapted to different study areas.

The Regional Ecology Branch has the mission to "develop tools for quantitatively describing the condition of ecological resources and their response to stresses at watershed and regional scales... Research involves regional surveys, process studies, and predictive modeling of the nation's waters, watersheds, and wetlands" (http://www.epa.gov/wed/pages/organization.htm). The research proposed here will develop statistical and simulation modeling approaches for looking at the status and response to stress of fish species distributions.

The Landscape Team has proposed a research strategy that focuses on "the long-term viability of native fish, amphibian, bird, and mammal populations that live within or depend upon wetlands, streams, rivers, and their associated riparian areas" (Baker et al. in preparation). The specific goal of this strategy is "to develop analytical approaches for prioritizing wetland and stream/river corridor restoration and protection to ensure the long-term viability of aquatic and aquatic-dependent biota." The research proposed here will look at the effects of natural and anthropogenic factors on fish species diversity in existing and modeled river networks. By focusing on large area studies this research hopes to provide a spatial context for the more detailed work performed by other members of the team. As the team continues its work in the Willamette River basin of Oregon, this research will suggest, through empirical analysis, how the fish biogeography of this basin is similar to, or different from, other basins in the west coast states. Simulation modeling of fish species distributions may help to identify particular network configurations of stream habitat and temperature regimes that promote fish species diversity.

1.5 Fish species as a unit of study

In view of the ESA's inclusion of distinct populations of species as protectable under the act, there is reason to have a goal of using evolutionary significant units, or stocks or runs to the extent they qualify, for the study proposed here. While the work on northwest salmon (Waples 1995) and a few other taxa, for example coregonid salmonids in the Great Lakes (Phillips and Ehlinger 1995) could possibly support this approach, in general there are few fish taxa for whom the necessary systematics work has been carried work. Furthermore, there is active research and debate in the ichthyology community about how and with which taxa to attempt to define sub-species units (see papers in Nielson 1995).

From the perspective of a large area view of fish species diversity as proposed in this work, the focus on species will cover, at the least, the currently recognized biological breeding units, at the species level, of the fish fauna. To the extent that the proposed analyses suggest relationships about species and their environmental correlates, these relationships presumably would pertain to all populations in some average sense. Of course, with the identification of populations, more precise statements could be made that could differentiate intraspecific variation. In the modeling research, a reasonable assumption would be that for any network within certain size ranges, the outcomes pertain to the populations of a species occupying such a network. If the modeling is successful and expands to attempt networks of high order, then the issue of units of study may need to be revisited.

****

2 Empirical analysis of possible mechanisms associated with fish species diversity

In this part of the proposed work, several statistical methods will be applied to data representing current known distributions of fish species in three west coast states and five east coast states, plus data representing possible factors influencing these distributions. The questions that this empirical work will address are:

- What environmental factors are associated with contemporary fish diversity and how do these relationships vary geographically?
- What factors are associated with human disturbance or alterations and how important are they?
- Do fish species group into distinctive assemblages when analyzed by their geographic distributions?
- What environmental factors are associated with any evident species assemblages?

In this section of the research plan, the rationale, theoretical framework, data, and methods for this study are discussed.

2.1 Why another study?

There have, indeed, been a number of studies in recent years looking at factors influencing fish diversity (see references above in section 1.1). These studies are far from conclusive, however, as rarely have even two studies been in complete agreement on reasons for observed distributions. Some studies of non-fish taxa, especially Currie and Paquin (1987) and Currie (1991) have argued support for a previously suggested hypothesis (in this case, Wright 1983), but contradictory results have been published later (Kerr and Packer 1997). Further arguments were presented by Guégan et al. (1998), with yet different results for fish.

Of the many papers referenced in section 1.1 on fish diversity studies, only four of these (Legendre and Legendre 1984, McAllister et al. 1986, Mandrak 1995, Rathert et al. in press) used representations of fish diversity on a regular grid covering the study area. (McAllister et al. 1994 used a regular grid to study coral reef fishes, but with very sparse occupancy; two Whittier papers, 1997 and in press, used sample data determined, in part, by a regular grid, but the analyses are by site.) The other studies all used a number of site samples for a set of river or stream networks, or lakes, that the samples were to represent (except for Kelso and Minns 1996 who used “areas of concern”, usually watersheds but sometimes including nearby lake waters). Thus variations in size of aquatic system are a factor in these studies. Of course, studies on a regular areal grid do not avoid this issue, since magnitude of aquatic system and accompanying habitat, regardless of how measured, will vary across equal samples of land area. However, representations of some potential correlates to fish response, such as climate variables and landscape patterns obtained from remote sensing, may be more appropriate on an equal area terrestrial grid. It would be highly desirable to represent fish diversity comprehensively over the river and stream segments and lakes that the fish inhabit, however this level of knowledge about current distributions is not available and estimates of such distributions would have, in general, a much larger degree of uncertainty than the area grid estimates.

Many studies have small sample sizes: of the 24 site level studies cited above in section 1.1 and Table 1, the minimum number of sites was seven, the maximum 2931, and the median 108.5; of the four grid studies, the minimum number of cells was 80, the maximum 907, and the median 332. (For site studies, all samples along all rivers or streams were included.) For the proposed work, the west coast states have 1261 grid cells with data, and the east coast states have 557 cells. Thus, the proposed work has more geographic detail (in total extent, or spatial resolution, or both) than many previous studies. The addition of EMAP data analyses could expand the scope more, including the approximately 600 sites sampled in the east coast sites, and some of the approximately 800 sites that may be sampled in the western US.

Two other aspects of the proposed work recommend themselves. First, by performing the same analyses on data from two different regions of the US, different lessons may be learned. Fish species distributions for east and west of the continental divide in the Rocky Mountains, for example, (McAllister et al. 1986) are known to be quite different in species numbers and composition. It would not be surprising if different environmental associations for the corresponding distributions are suggested by analysis. Second, by comparing the grid cell data with site data collected by EMAP, another dimension of the distributions can be investigated.

We will be able to look at how environmental associations differ between these data and how observed species pools compare with estimated.

We believe that this study will be the most detailed study of this type, in terms of the combination of spatial extent and resolution, in terms of the comprehensiveness of the fish fauna, and in terms of the comprehensiveness of the representation of environmental factors, including those associated with human disturbance.

2.2 Hypotheses about the distribution of diversity

A useful classification about the many hypotheses of geographical variation in species diversity is in Brown and Lomolino (1998), where effects of space and time are contrasted with effects of productivity, abiotic stress, and biotic interactions. Huston (1994) and Rosenzweig (1995) have comprehensive discussions of patterns of species diversity and mechanisms influencing these patterns. Matthews (1998) discusses many issues in the geographical patterns of fish species diversity. Hypotheses about species diversity are not all mutually exclusive, and vary in importance among places and species. For this study species diversity will include both species richness, or numbers of species, and composition, although most effort will be focused on richness. Some questions expressing these hypotheses are:

- Does species richness increase with amount of area (water) sampled?
- Does species richness increase with amount of available energy, decreasing latitude, or decreasing elevation?
- How is species richness affected by biological interactions?
- Is species richness constrained by thresholds in tolerance to physical conditions?
- Does species richness decrease with disturbance, or decrease after initially increasing?
- Does species richness increase with stable temporal patterns in physical conditions?
- Does species richness increase with increased variability of habitat?
- Does species richness increase with decreasing physical limits to dispersal?
- Is species richness constrained by the contingent effects of geological history, or other local factors?
- Is the relationship between species richness and possible explanatory factors different than the relationship between higher taxon richness and such factors?
- Is the relationship between species richness and possible explanatory factors affected by the spatial scale of analysis?

Data to represent these hypotheses will be discussed below under explanatory variables.

2.3 Data to be used in analysis

Geographical extent

The geography for these studies will be the three west coast states of the US (Washington, Oregon, California) and five east coast states in the Chesapeake Bay area (Pennsylvania, Delaware, Maryland, Virginia, West Virginia). This extent is determined by agreements between the EPA and TNC (Master 1994) that specified these states to be included in TNC's data preparation efforts. This study region is divided into a regular grid of cells approximately 650 km^2 in size using EMAP grid methods (White et al. 1992). These grid cells will be the units of analysis. In studying species distributions at these scales, decisions about the relative importance of precision versus accuracy are necessarily made. Smaller analysis units provide more precision but, for these kinds of studies, often at the expense of accuracy (White et al. 1999).

Taxonomic scope

This study will focus on all native non-marine fish species that are resident in the region of study for some part of the year, including lampreys (Master 1994). The authority for inclusion or exclusion of species in these data is TNC, in particular their Vertebrate Characterization Abstract (VCA) (Master 1994). The VCA is a dynamic database receiving continual updates, so the data used here will reflect the status of this database at the time the data were prepared and delivered to EPA, plus any updates forwarded by TNC (Master personal communication).

Response variables

Species are classified by TNC according to occupancy status as absent, possibly present, probably present, or confirmed for each grid cell. We will convert this classification to only presence/absence by recording either probable or confirmed as present, on recommendation from TNC (Master personal communication). Species are further classified by origin as native or introduced. We will separate species into native and exotic (introduced) so that either or both can be used as response variables.

TNC has reported that additional quality controls will be performed on these data between March and July 1999 (Master personal communication). We will update our data when this work is completed. A sample of the distribution maps are in Figure 1.

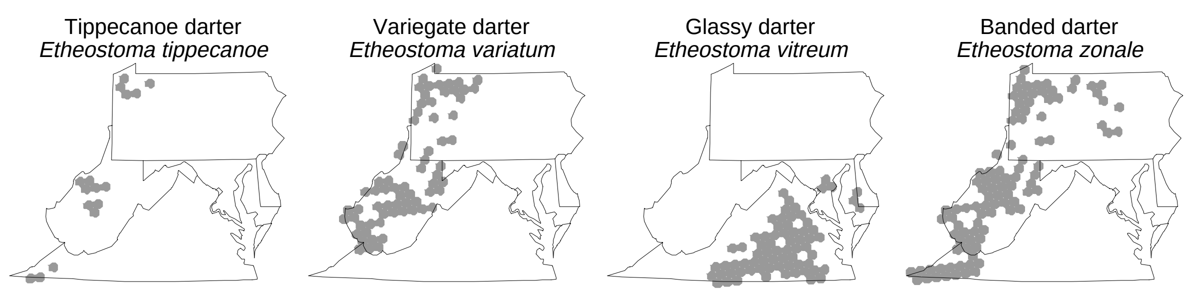


Figure 1. Four sample species distributions in middle Atlantic states.

In addition to investigating the richness of all native species as a response variable, subset effects and assemblage composition will be studied. One type of subset that will be considered are species with TNC conservation status of G1, G2, or G3, that is, species that are considered by TNC to be “critically imperiled”, “imperiled”, or “vulnerable” globally. G3 species typically have between 21 to 100 occurrences worldwide with between 3,000 and 10,000 individuals; G2 and G1 species have even fewer occurrences and individuals (Master 1991).

Another type of subset could be species with similar life history traits, or, alternatively, a systematic sample of species in life history space. With respect to life history space, the intention would be to place all species in a conceptual structure such as a three dimensional space of variation where the dimensions could be size of organism, movement ability, and habitat utilization (Schumaker, personal communication), or perhaps other factors relevant to fish. In studying the sample from life history space, the objective will be to investigate how the sample performs compared to the full population; collaboration with other members of the Landscape Team in formulating and conducting these studies will be important.

Explanatory variables

The development of explanatory variables will be driven by hypotheses about the distribution of fish species diversity. Before discussing hypotheses individually, data available from previous studies is noted. From earlier work (O'Connor et al. 1996, White et al. 1999, Rathert et al. in press), data are available for the region of study on January minimum temperatures, July maximum temperatures, annual precipitation, elevation, and variables that were a part of the Loveland et al. (1991) seasonal land cover classification. All these data, plus other variables derived from them, were developed on a grid of 1 km^2 pixels across the conterminous US.

A major research effort in the Chesapeake Bay region (same definition of study area as this work proposes) has been undertaken by Jones et al. (1997) to assemble and correlate a large number of indicators of watershed condition based on landscape-level data. The work included a comparative assessment of nine of the indicators for all eight-digit hydrological accounting units (Seaber et al. 1987) in the study area. The indicators developed in Jones et al. apply to a number of the hypotheses to be considered here and these data will be evaluated and incorporated as appropriate. In addition, a primary source of data for Jones et al. was the Multi-Resolution Land Characterization (MRLC), a Thematic Mapper based 30 meter resolution dataset that will eventually cover the entire conterminous US (http://www.epa.gov/mrlcpage/text.html). As west coast data from MRLC become available, they will be incorporated as well.

Species-area. Many studies support this hypothesis, an outcome predicted by the explicit relationships of Wilson and MacArthur (1967). An analysis of aquatic systems on equal area terrestrial grids does not provide a control for this factor as it could with terrestrial species. This study will develop a representation of water volume by grid cell in several ways. One method is from a function of lengths of streams of different orders plus lake areas. Another way is from estimates of runoff using, for example, methods described by Bishop et al. (1998). An additional idea is the drainage area of large basins that include each cell (Bradford personal communication).

Species-energy. Wright's 1983 study formulated for analysis and tested the energy hypothesis; many subsequent studies have supported this idea. A large number studies have looked at latitudinal gradients in richness as possible results of this or other mechanisms (see Brown and Lomolino 1998, for a summary); some studies have focused on elevational gradients in richness (see Rahbek 1995, for discussion). Some studies have used potential and actual evapotranspiration as indicators of available energy. Others have used net primary productivity, integrated received radiation from remote sensing data, solar radiation, and mean temperature.

For aquatic system response, one type of ideal variable would be aquatic system productivity, however, these data are not available in general.

For the species-energy hypothesis, this study will compare data from Loveland et al. (1991) on total Normalized Difference Vegetation Index (NDVI), a measure of terrestrial vegetation productivity with other estimates of actual or potential evapotranspiration (for example, USSR National Committee for the International Hydrological Decade 1977). In addition, work of Reed et al. (1994) on measures of vegetation phenology using NDVI may provide variables for investigating this hypothesis as well as the habitat heterogeneity hypothesis.

Biological interactions. Competition, predation, and other interactions are difficult to represent with empirical data at the scales of studies reported here. Many theoretical and observational studies, of course, show effects on local assemblages of these ecological processes. Some investigations have looked at possible saturation effects from species-area curves. In some simulation modeling studies, concepts of local carrying capacity and species replacement have been implemented. Although the scale of investigation may preclude effective analysis of these effects, this study will consider possible ways to represent them. These effects may be better studied by the simulation modeling proposed in the second half of this research plan.

Environmental-physiological thresholds. From animal (and plant) physiology research, most species are known to have limits to their distributions based on temperature, water availability, oxygen, and other chemicals acting either as nutrients, toxins, or both. These factors are often confounded. Many studies have used extreme temperatures, precipitation totals, or sampled values of chemicals to represent these effects. For this hypothesis, this study will consider redeveloping the existing temperature data described above to obtain minimums of daily minimums rather than minimums of daily means. The work here will also investigate the modeling of water temperatures from sparse networks of water temperature observations plus correlations with elevation or air temperature (itself modeled from elevation - see Marks et al. 1990). Other effects related to water availability using hydrological regime and precipitation measures will be investigated.

Environmental disturbance. The intermediate disturbance hypothesis provides an interesting source of conjecture about the distribution of diversity (see Rosenzweig and Abramsky 1993 for a discussion). However, definitions of “disturbance” and “intermediate” may vary greatly from that observed in Connell’s (1978) original work in order to formulate a test for any given study. Huston (1994) pointed out that a definition of disturbance related to species richness necessarily results in a circular argument. Moyle and Light (1996a, 1996b) suggested a number of patterns that invasions of exotics may have on existing fish assemblages. In addition to exotics, the disturbance hypothesis could include other indicators of possible human influence such as density of human population and roads, for example, or land use patterns. In studying disturbance with these types of data, a space-for-time tradeoff would be asserted insofar as the intermediate disturbance hypothesis, for example, relates to the time frequency of disturbance.

For this hypotheses, this study will extend the data developed by Wickham et al. (1997) for Oregon on human population and road density. These data can be derived from data available from the USGS. The human use index, the proportion of land in urban or agricultural uses, as described in Jones et al. (1997), will be applied as well. For effects on fish, road crossings of streams or length of roads adjacent to streams may be as or more important than road density (see Jones et al. 1997 for examples of these variables). In addition we will have data on presence of exotic fish species from TNC. Other ideas to be investigated are influences of agriculture, forestry, grazing, and mining on the fish response. Although past work in the Biodiversity Research Consortium (including Wickham et al. 1997) has attempted to represent forestry and grazing disturbance factors, no satisfactory representation has been developed.

Environmental stability. Although the stability-diversity hypothesis has been largely repudiated as a widely applicable principle (see Shrader-Frechette and McCoy 1993, for one analysis of this issue), at least one study of those reviewed (Fraser and Currie 1996) investigated environmental stability represented as temporal variability in long-term climate and salinity data. For this hypothesis, this study will revisit existing data to add variables representing temporal variability in air temperature and precipitation. The earlier work by Daly et al. (1994) used previously has been supplemented by a number of additional variables now available through the Oregon Climate Service web site (http://www.oce.osu.edu/prism/prism_products.html).

Environmental-habitat heterogeneity. Landscape ecology has argued that increased habitat heterogeneity, or variety of local habitats, enhances potential species richness (Forman and Godron 1986). Some investigators (e.g., Rosenzweig 1995, Davidowitz and Rosenzweig 1998) argue, conversely, that cause and effect are reversed: more species (due to some other mechanism) create more habitats. Studies have used variation in number of habitat classes, elevation, climate variables, or landscape metrics to investigate this hypothesis. For this hypothesis, this study will calculate spatial variation within analysis cells for many of the variables mentioned above as possible measures. Another variable that will be investigated is the number of different stream orders and whether lakes are present in a cell (Baker personal communication). In addition, we will investigate spatial variation in the climate and soil spatial clusters of Hargrove (http://www.esd.ornl.gov/~hnw/esri). Even so, none of these variables may capture the type of local heterogeneity that influences fish diversity (see Matthews 1998).

Limits to dispersal. Besides physiological limits to species distributions, physical abiotic limits occur, for example, at the terminal points in river or stream networks for most fish. The counterpart effect for terrestrial organisms would be the “peninsula” effect (Simpson 1964, Brown and Lomolino 1998). For fish, an explanation using physical limits would often probably be confounded with physiological limits (cold water), species-area or species-energy (low levels), or other factors. For limits to dispersal, this study will develop one or more variables derived from stream orders determined for the national stream network databases (http://nhd.fgdc.gov; http://www.epa.gov/owow/nps/rf). One example could be proportion of first order streams by cell. Another idea could be number of adjacent cells in associated networks (Bradford personal communication). This hypothesis is also linked to historical effects of these limits and their breeching during glacial warming periods, for example (Matthews 1998), as discussed in the next hypothesis.

Historical and local effects. Environmental history over geological time may affect diversity patterns. Smith (1981) and Matthews (1998) discuss the great effects on fish distributions of Pleistocene glaciation in North America and the role of the lower Mississippi River system, for example, as a major refuge. Some studies have represented these effects of Pleistocene glaciation, for example McAllister et al. (1986), by coding sites (or grid cells) with a binary indicator of glaciation history. In the western study area, northern and high mountainous areas are affected and in the eastern study area the northern most areas are affected.

Glaciation and other historical effects like river capture will be represented by coding cells by their location within major drainage systems, present or past, that may have influenced colonization by fish fauna (Rathert et al. in press). A related idea, also explored by Rathert et al., is distance to possible sources of species from oceanic migration. McAllister (personal communication) has suggested investigating the ratio of euryhaline to primary freshwater species as another way to look at ocean effects.

Higher taxon diversity. McAllister (personal communication) has suggested comparing species diversity relationships with those of higher taxon diversity. Gaston and Williams (1993, Williams and Gaston 1994) have explored this issue with birds, butterflies, bats, and ferns.

Spatial scale effects. Several studies show different factors operating at mesoscales than have been shown or hypothesized to operate at macroscales (Fraser 1998; Rathert et al. in press; White et al. 1999). Meentemeyer and Box (1987) provide a general discussion and review of effects of spatial scale on analyses and of tools for analyzing scale effects. Withers and Meentemeyer (1999) and Fortin (1999) have updated these discussions. In addition to methods described below, in particular, regression trees, that help to analyze scale effects, this study will consider other methods to explore these issues.

One set of questions about spatial scale contrasts local with regional species richness. Is species richness over large areas purely the result of the aggregation of the effects of local processes, or is local richness a linear function of regional richness? Are local fish assemblages saturated in species relative to regional pools? These questions attempt to disentangle the relative influences of local effects of physical conditions and ecological interactions versus regional biogeographical influences on the regional species pools. Ricklefs (1987) has a summary of these issues; Matthews (1998) discusses effects on fish assemblages; and Griffiths (1997) is a recent analysis for fish in North American lake assemblages. A challenge in these studies is defining local and regional scales. For the TNC data, it will be possible to compute a large number of species-area curves from single grid cells to various aggregations of these. Also, the statistical methods described below can be used on aggregated cells as analysis units as well as on the individual cells.

Another issue raised here is how to define regions. Omernik and Bailey (1997) argue that ecological regions, or "ecoregions", better represent the spatial structure of ecological processes than watersheds. The recent World Wildlife Fund sponsored project takes a different view (Abell et al. 1999). Matthews (1998) feels that in most cases river basins will be a better spatial unit with which to study fish assemblages. By grouping grid cells into both of these types of spatial units, environmental relationships determined in both ways can be compared. In addition, through cluster analysis (e.g., Hughes et al. 1987) and boundary detection methods (e.g., Fortin and Drapeau 1995), the empirical spatial structure of assemblages can be compared with both of these types of spatial units.

2.4 Analysis methods

The analysis strategy of this work recognizes that explanations for the distribution of fish diversity are likely to suggest multiple rather than single mechanisms, to have patterns that are contingent, at least in part, on the history and geography of the study region, and to have a hierarchical structure whereby different mechanisms operate at different scales (O'Neill et al. 1986, Pickett et al. 1994). To better assess different mechanisms at various scales, we plan to use several different analysis methods that together may help to reveal this complexity.

Multiple linear regression. Regression analysis is widely used and there are a large body of techniques for improving the understanding of a problem through regression. Variable selection techniques, transformations of variables, and analysis of residuals are some of the techniques that will be used in this work. Regression attempts to fit to a response a linear combination of predictors for a "global" explanation of associations. In Rathert et al. (in press), after initial investigation, the study region was divided into two part and each subregion fit with separate regressions. The initial investigation was aided by work with regression trees.

Jager (personal communication) has suggested using null models including and excluding predictors representing individual hypotheses to distinguish between possible alternative explanations that may arise from multiple linear regressions considering all possible predictors. Similarly, Ferraro (personal communication) suggested "hierarchical" regression in which individual predictors are introduced into the model in the a priori order in which they are hypothesized to have logical or causal precedence. These suggestions will be investigated.

Regression trees. This technique subdivides the observation space in a hierarchical manner, choosing the single value of one of the predictors that minimizes residual variation at each point of division (or split). Splitting continues until stopping criteria have been satisfied. The hierarchical nature of the results can reveal scale relationships in associations. We have considerable experience with this technique (O'Connor et al. 1996, White et al. 1999, Rathert et al. in press, Sifneos et al. submitted). An example of output from a regression tree is in Figure 2.

Neural networks. Guégan et al. (1998) used artificial neural networks to predict fish species richness in 183 river systems across the world from predictors representing river discharge, net primary productivity, and drainage area. Like regression trees, neural networks allow for non-linear associations between predictors and response. The fitting is done with a set of simple weighted linear functions between the predictors and an intermediate ("hidden" in neural network terminology) set of state variables, and then a second set of similar weighted linear functions from the intermediate states to the response. The methodology attempts to determine the optimal set of weights, given the number of intermediate states as set by the investigator, to minimize residual variation in the fit (Edwards and Morse 1995, Lek et al. 1996).

Spatial autocorrelation. In neural networks, as well as in multiple linear regression and regression trees described above, spatial autocorrelation may complicate fitting (Haining 1990, Venables and Ripley 1997). A locally developed method for handling spatial autocorrelation in

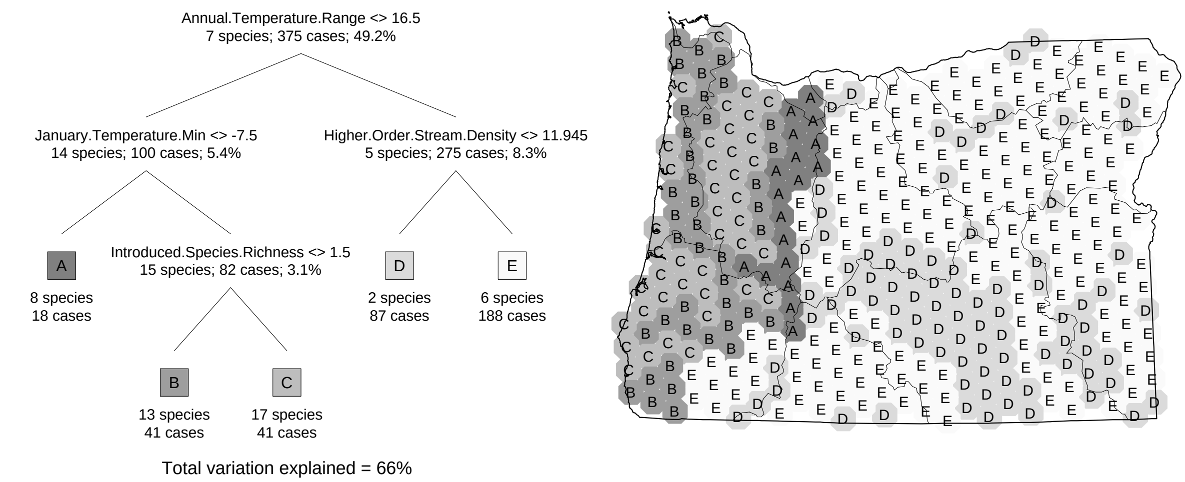


Figure 2. Example of regression tree diagram and corresponding map of observations in terminal nodes. The response variable is native fish species richness. In this case a different predictor variable prompted each split. Adapted from Rathert et al. (in press).

regression trees (Urquhart personal communication) will be investigated, and Haining (1990) discusses procedures for treating this effect in regression studies.

Model validation. The implementations of multiple linear regression, regression trees, and neural networks that will be used (Chambers and Hastie 1992, Venables and Ripley 1997) include internal validation using cross-validation or withheld subset prediction.

Analysis of assemblages. A number of techniques have been developed in community ecology to study the structure of species and sites and their relationships with environmental correlates. A recent comprehensive study of this type with fish data is Angermeier and Winston (1999) where canonical correspondence analysis, detrended correspondence analysis, multi-response permutation procedures, principal coordinates analysis, redundancy analysis, and spatial trend surface analysis were used. An example of a biogeography application is Whitehead et al. (1992); a general discussion about applications in biogeography is in Hengeveld (1990); and general advice and cautions with multivariate analyses are in James and McCulloch (1990). Non-metric conceptual clustering (Matthews et al. 1991) is another method for grouping. In addition to multivariate approaches, Billheimer et al. (submitted) developed the reduction of proportional compositional data to a univariate response using the norm of an additive logratio transform of the proportions. This statistic provides a measure of evenness of the compositions. Other approaches for reducing compositions to a univariate response are summing class ranks which provided a magnitude measure, or principal component measures. The analysis of assemblages part of the work will be developed in collaboration with statisticians in the landscape team and elsewhere.

2.5 Comparison with EMAP data

EMAP has conducted sampling of streams in the middle Atlantic highlands region during the years 1993-1998. Approximately 600 sites in total were sampled, not all in every year. Among the data collected were presence of fish species, as part of the development of indices of fish assemblage integrity (McCormick et al. submitted). This data set offers the opportunity to compare associations determined from TNC data with confirmations or alternatives determined from EMAP data. In addition species lists can be compared, since it is likely that, on the one hand, EMAP will not have sampled all species estimated to be present, and, on the other, EMAP may have sampled species not previously known to be present (e.g., Whittier and Hartel 1997).

In Oregon, EMAP has been collecting samples during the years 1992-1998 in pilot studies and a total of approximately 250 sites were sampled in comprehensive pilots in 1997 and 1998, not all in both years. This work also supported the development of fish assemblage integrity indices (Hughes et al. 1998). EMAP is now planning a comprehensive assessment in 12 western states, called the “Western Pilot” (http://www.epa.gov/emap/wpilot), during which approximately 600 sites will be sampled. The proposed empirical investigations will use these data as they become available for the three west coast states.

For purposes of comparing species lists, EMAP data will be allocated to the grid cell structure of TNC data. For purposes of association analysis, EMAP data will be treated as sites rather than allocated to grid cells. Consultation and collaboration with EMAP scientists will help to determine how to assemble predictor variables from data collected by EMAP and prepared by this project. A major source of potential explanatory variables is expected from the landscape component of EMAP, both from existing work on the east coast (Jones et al. 1997) and corresponding efforts in the Western Pilot. Similar methods of analysis will be used for both TNC and EMAP data, except where EMAP sample sizes may be small. Regression trees are more constrained by sample size than other methods, performing more usefully as sample size increases beyond 100 cases.

****

3 Simulation of fish species distributions

One side effect of the focus on single species by the Endangered Species Act has been the development of methods to assess the viability of the populations of a species. A number of computer models have been developed that simulate the movement and breeding of individual organisms on a landscape over lengths of time sufficient to measure persistence. Some models have been generalized to study any species that has appropriate behavioral dynamics. Research models (e.g., Schumaker 1998) and commercial models (e.g., http://www.ramas.com/software.htm) are available.

No comparable effort has been made to look at the dynamics of species assemblages and diversity. Although a number of research models have been published (see above in section 1.1), this work usually has been a one-time effort, rather than a generalizable approach. Furthermore, few if any organismic models, including spatial demographic models, have been implemented in the setting of a directed acyclic network to simulate population processes in river systems.

The research proposed here will explicitly simulate the dynamics of a hypothetical pool of mobile species occupying a network, where sub-populations of a species may be colonizing new segments if their attributes match the environmental attributes of the segment, possibly replacing existing species if their match is better, possibly suffering local extinction, possibly expanding in size, and possibly migrating to neighboring segments. The objects of analysis will be sub-populations, however, rather than individual organisms. Sub-populations will be defined as all individuals of a species occupying a segment, or link, of the network.

This work will be exploratory. The initial work will be to develop a prototype of a model with an abstract network and a hypothetical pool of species, and to tune parameters to see if reasonable behavior results. If promising, the work will evolve into the use of actual networks and species pools. The questions asked by this work are:

What are the important aspects of species and the environment needed to simulate the distributions of assemblages and diversity over time?
What are the sensitivities of various parameters in the behavior of the simulations?
How well does the behavior of the model match contemporary distributions?
How do the effects of simulated human disturbance, as in introduction of species, or reduction of habitat, or placement of dams, affect results?

The simulation modeling differs from the empirical analysis in the types and spatial scale of analysis units, and in the hypotheses to be investigated. The analysis units for the simulations will be network links rather than hexagonal sampling areas used for the empirical analyses, and the hexagonal areas are large enough to include a number of links in most types of river networks. Thus there is a spatial smoothing effect in the empirical analyses relative to the simulations. The hypotheses to be studied in the simulation modeling will include, at least initially, biological interactions, physiological thresholds, environmental disturbance, environmental-habitat heterogeneity, and local versus regional richness spatial scale effects. Some hypotheses, for example biological interactions, will be likely to be better studied in the simulation framework. Correspondingly, the empirical association analyses may provide evidence of the importance of certain hypotheses at certain spatial scales. Such evidence may be useful in considering additional parameters and relationships for the simulation modeling.

3.1 Representation of the stream network

The stream network will be conceptualized as a directed acyclic graph of nodes and links (Foulds 1992), where each node has two or more, but usually only two, upstream links entering, and one or more, but usually one, downstream link leaving. The nodes will not carry any information relevant to the problem; they will only serve as conceptual connections. The links, however, will be the local units of the environment and will have a set of properties whose values are subsets of the total environment represented over the network.

Each link will have one or more attributes derived from the network topology. Haggett and Chorley (1969) review a number of possible topological measures; of these, Strahler and Shreve orders (Figure 3; Chorley 1969) are two estimates of the “size” of links. Strahler order is a measure of complexity of branching while Shreve order is the number of first order (headwater) links upstream. Strahler order is more commonly used, but Shreve order may better represent the magnitude of hydrological properties (Osborne and Wiley 1992). Another possible measure is that of Horsfield (Horsfield and Woldenberg 1989).

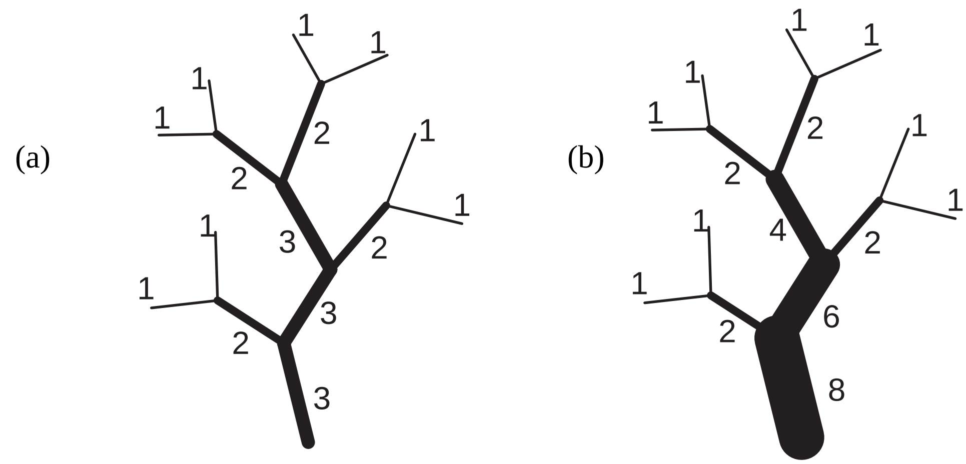


Figure 3. Strahler (a) and Shreve (b) stream ordering systems. If *a* and *b* are upstream links that merge to form *c*, then the order functions, ω, are:

Strahler: $\omega_c = \text{MAX}\ [\ (\omega_a + \omega_b)\ \text{DIV}\ 2 + 1, \text{MAX}\ [\omega_a, \omega_b]\ ]$

Shreve: $\omega_c = \omega_a + \omega_b$

where MAX is the function returning the largest of its arguments, DIV is integer division (truncation), and headwater links are assigned order 1 in either system.

Hypothetical stream topologies for developing the model can be generated with recursive functions using a simple parameterization to obtain desired branching properties and sizes. The “Q model” (Costa-Cabral and Burges 1997) is one such method that incorporates a single parameter that is the ratio of the probability of internal branching to the sum of the probabilities of internal plus external branching. In the early stages of development of the model networks will be generated stochastically according to these probabilities.

The environment of a link in the network

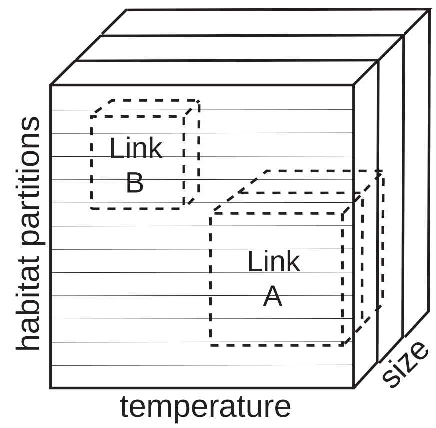
Ward (1998) describes a conceptual model of the complexity of the environment of riverine landscapes. Abstracting from this complexity, in this research the properties of the environment represented in the network will initially be in three dimensions: temperature (of the water), habitat partitions, and size. Temperature and habitat can be conceptualized as a two dimensional space where temperature has a continuous metric on its dimension and habitat has an integral number of discrete partitions. Habitat is taken to mean all components of fish living requirements other than temperature. The partitions could be taken to represent niches, but will be fixed as a property of the environment and will not be modified by populating species. Any link contains a contiguous subset of this two dimensional space, conceptually shaped like a box (Figure 4).

A link’s subset of the temperature dimension will be determined by the position of the link in the network. In general, links farther from the exit will have lower temperatures than those closer.

Temperature as a more complex function of elevation, land cover, and size of link will be developed during the course of research to simulate more realistic conditions.

A link's habitat subset will be a range on the habitat scale that will include one or more partitions. The subset range will be determined randomly with modification to enforce some degree of spatial autocorrelation. That is, neighboring links will be different, but not too much, from each other. Because this range will be determined on a continuous scale, the partitions at the upper or lower end of the subset may be partial.

Figure 4. The environment for a network is a three dimensional space varying temperature, habitat (divided into partitions) and size (a function of stream order). A link in the network contains a subset of the total environment.



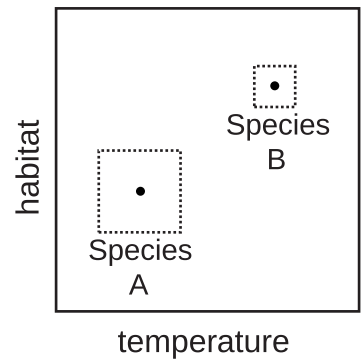
The size of a link's environment will be a function of the stream order. The size will be represented as an integral multiplier on the number of partition spaces available in the link (Figure 4). For example, first order streams would have only one set of partitions, second order would have two, and so forth. Extra spaces will be used for growth of a sub-population and for occupancy by a sub-population displaced by a more fit species. So each partition space, defined as one size unit of a partition, can have only one sub-population.

3.2 Representation of fish species and sub-populations

There will be a (initially hypothetical) pool of species, each of which has temperature and habitat requirements that are a subset of the total environment available in the network. A regional pool of species for basins appears to be a reasonable starting point for investigating patterns over a basin (Matthews 1998). Species requirements can be conceptualized similarly to the environment of the network. Each species will have a range of contiguous temperatures and a range of contiguous habitat within which it can exist. The combination of the two ranges creates a box of requirements (Figure 5). The center of the box is the optimal value for the species' existence. Species requirements will be random samples, both in size and location, from specified distributions, initially uniform. Although many fish species are considered generalists, at least in feeding behavior (Gerking 1994), this conceptualization of fish living requirements will allow a range of requirements along a generalization-specialization continuum.

Each species also will have three other parameters: (1) mobility range, *m*, specifying how many links its sub-populations may move in one unit of time, a parameter randomly drawn from an exponentially decreasing distribution; (2) rate of population growth, *r*, the intrinsic rate of natural

Figure 5. The requirements for a species are a subset of the total environment.



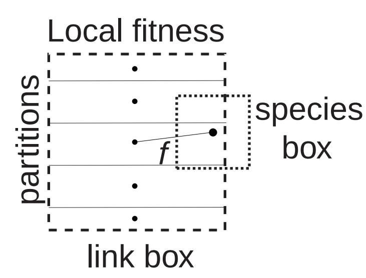
increase (Pianka 1978), randomly drawn from a lognormal distribution; and (3) extinction rate, a probability of local extinction. The extinction rate will be a function of *r* and the size of the habitat and temperature ranges.

The unit of a species that occupies one link will be called a sub-population. Initially, all sub-populations of a species will have the same parameters; this specification could be relaxed as the modeling evolves to simulate local adaptation or other types of intraspecific variation.

3.3 Local fitness

When a sub-population of a species moves into a link, its ability to colonize will be measured by what is called here local fitness. First, the center of the species' requirements box must lie within the link's environment box; if not, the species cannot be present. The distance from the center point of the species' requirements box to the center of the habitat partition in which the species' center is located, *f*, is taken to be a measure of local fitness with lower values corresponding to greater fitness (Figure 6). This model of local competition attempts to encapsulate a wide variety of competitive behavior (Wootton 1990).

Figure 6. Local fitness is the distance from the center of the species' requirements box to the center of the habitat niche box within which the species center is located.



When a sub-population of a species is migrating, but before it attempts colonization, it may pass through a link if its requirements box intersects, in any amount, the link environment box. Thus a species may migrate through parts of the network that it cannot colonize.

3.4 Simulation of biogeographic and ecological processes

Time units and initial conditions

This modeling will attempt to simulate the colonization of a river network by a pool of species and the ensuing vicissitudes of species distributions due to their own interactions, possible changes in the pool of species, changes in environmental conditions, or changes in the topology of the network. The time frame for these kinds of effects would be centuries or millennia, long enough for some sorting out of species distributions, but not long enough for speciation to significantly affect the species pool. The time steps in the simulation will then be considered to be units of between one and perhaps ten years. Yearly migrations of diadromous species will not be simulated, at least initially.

The simulations will start under one or more alternative sets of initial conditions. To simulate the occupancy of a river network following a major change in environmental conditions, such as retreat of glaciation for example, species will enter at the mouth of the network, their probability of entering in a time step being proportional to their mobility parameter. Under this type of initial condition, the environment of temperature and habitat may be gradually altered over some beginning period of time to simulate initial colonization by cold water species that in later periods would be excluded from entry, for example. Alternative initial condition could be a random placement of species in the network, or current distributions in a real network.

Colonizing

A sub-population that has arrived at a link can remain as inhabitants if the species optimal requirement (center point of requirements box) is within the link environment box, and, if there is an empty habitat partition within the requirements box. If there is more than one partition available, then the sub-population will occupy the one that maximizes local fitness. If there is already a sub-population of the species resident in the link, the new sub-population will not colonize.

Replacing and relocating

A sub-population may replace a sub-population of a different species occupying a habitat partition. This can happen at any time step if the local fitness of a species is greater (smaller f value) than the current occupant. The replacement function will be a weighting of the difference in local fitness, such that small differences will not usually result in replacement, simulating a “priority effect” (Matthews 1998), with a stochastic component applied to the weighted difference. A sub-population that is replaced may relocate to another size space in the same partition, or to another partition if its species requirements box intersects the partition and the partition is available. The closest such partition will be occupied.

Dying out

If a species is replaced and no other partitions are available, its sub-population in that link becomes locally extinct. In addition, a stochastic effect will be implemented whereby a sub-population can become locally extinct if its extinction probability, *e*, times its fitness distance, *f*, exceeds a randomly determined threshold.

Growing

A sub-population may grow to occupy more size spaces in the same habitat partition if spaces are available, and, if its growth rate, *r*, divided by its fitness distance, *f*, exceeds a randomly determined threshold. Growth will only occur in the partition currently occupied. Under this restriction, species only occupy one partition, but several species can occupy the same partition.

Migrating

A species may migrate to neighboring links if the links are within the mobility range of the species, *m*, and the links' environment boxes intersect the species requirements box. A species may remain in migration state until the next time step when it must either colonize or die out in a link.

3.5 Experiments

The types of experiments to be conducted with this simulation model can be categorized as sensitivity studies, effects of human alterations, and extensions of the model.

Sensitivity studies

The currently proposed parameterization of the model has five parameters for each species, two parameters for each network link, at least three parameters governing replacement, growing, and local extinction of sub-populations in links, plus parameters related to the generation of the network size and shape (for example, the "Q" parameter) and the size of the species pool. It will be important to understand the behavior of these parameters and how sensitive possible outcomes are to them.

Drechsler (1998) has a useful discussion on studying the sensitivities of complex models, with particular attention to population biology models and models with non-linear dynamics. The model proposed in this research will certainly exhibit non-linear behavior due to the interactions of species with the topology of the network. Drechsler's "sensitivity analysis of sensitivity analyses" may help to focus the tuning and initial experimentation of the model in order to learn how it is behaving.

Another approach to model assessment is that of Reynolds and Ford (1999). Their "Pareto Optimal Model Assessment Cycle" uses genetic algorithms and related techniques to explore the parameter space of a model in the attempt to identify parameterizations that satisfy an optimal number of assessment criteria. Assessment criteria may be empirical data, theoretically derived criteria, or criteria determined in some other way. As a method for systematically investigating proposed model parameters against multiple measures of performance, these ideas have appeal for the work proposed here.

Effects of human alterations

Two major human effects that will be simulated are species introductions and habitat loss or alteration. Species introductions can be simulated by adding a species to the species pool and introducing a population either at the entrance to the network or to some random link. Exotics may often be at the generalist end of the habitat utilization spectrum, but not always. Experiments will look at whether exotic effects on diversity occur and what the effects are. Alteration of environmental conditions can take many forms. A degradation of habitat can be simulated by narrowing the range of habitat and temperature. Change in climatic conditions can be simulated by raising the range of available temperature. Habitat changes can be changed over the entire network, or only in portions to simulate more local effects.

Extensions to the model

Since many aspects of the spatial ecology of fish are dependent on life stages, the development of a species life stage interaction model could make this work more realistic. One concept for implementing this would be a species-life stage interaction matrix with entries for each species and life stage combination. Entries could indicate predation, consumption, competition, or no effects, for example. These processes would then be activated at the phases of colonization, replacement, growth, and so forth.

Two other areas of interest in extending the model, if it proves successful, are in representing lakes, first, and dams, second. Lakes would be conceptualized, initially, as links in the network with (1) a special habitat and temperature environment, and (2) constricted connections to neighboring links. Dams would be conceptualized as constrictions on link connections, or, as links where the environmental conditions make large discrete changes in habitat and temperature. Investigation of the effects of species requirements would probably accompany these extensions.

3.6 Implementation

Because this work is exploratory, the development of the model will be done in a computer modeling environment that enhances rapid prototyping above other capabilities. To this end, an interpreted language with expressive functionality and clean syntax will be preferred. Because of past experience, two candidates for this work will be S-Plus™ and Mathematica™. In addition, two newer languages, Python (http://www.python.org) and Haskell (http://www.haskell.org), that have some advantages in prototyping, especially in representation of data structures, will be considered. All of these development environments provide some form of object oriented problem representation.

****

4 Collaboration: team, branch, agency, TNC

An important aspect of the scientific and social institutionalization of the proposed research is the hierarchical collaboration that is envisioned. Within the landscape team there will be important collaborations with the agricultural riparian and fish response project (Van Sickle and Wigington) in sharing ideas and data on empirical modeling, and with the modeling group, who have already been of great help in formulating ideas for this proposal, in relating species diversity modeling to meta-population modeling (Schumaker and Heppell) and in relating diversity modeling to landscape systems modeling (Leibowitz). Within the regional ecology branch it is hoped that there will be possibility of collaboration with the EMAP surface waters group on many aspects of the empirical analyses, including sharing data and ideas and jointly conducting analyses. Also the input of the fisheries experts in EMAP surface waters in the simulation modeling will be highly valuable as well.

In the wider agency setting, it is hoped that there will be possible collaborations with EMAP fisheries experts at other divisions and laboratories and with the Regional Vulnerability Assessment project (http://www.epa.gov/nerlrvva/index.html), particularly in its efforts in the middle Atlantic region.

Because of the long association of the biodiversity research program in EPA with TNC, it is hoped that a continued collaboration with scientists at TNC will enhance use of their data and allow for sharing of ideas and possible joint analyses.

5 Projected work load and time frame

The work that is proposed here has a scope that can be accomplished largely by the principal investigator should funding amounts and mechanisms limit assistance. The time frame for these investigations is approximately three years. Without help, it is projected that at least one major and several minor publications can be prepared on each part of the project by the end of the project period. The scheduling of activities will be to advance on each part somewhat in parallel, as shown in the table below.

Schedule	Empirical Study	Simulation Study
Year One	Data preparation	Model development
Years Two and Three	Analysis and interpretation	Experiments, analysis, and interpretation

Clearly more can be accomplished if assistance is available. The first level of help would be with data preparation for the empirical study. For this work, skills in environmental data and geographical information processing will be most useful. Additional skills that could benefit empirical analyses would be statistical analyses of environmental data, especially multivariate data sets. For modeling work, help would not be needed until the work has evolved to where more efficient computer implementation would be useful, at which time computer programming skills would be essential.

6 Quality assurance

Response data to be used in the empirical analyses, from both TNC and EMAP, have been collected under EPA approved QA plans. A number of explanatory variables were developed under the EPA sponsored Biodiversity Research Consortium that also had an approved QA plan. A Quality Assurance Project Plan discussing these and other issues accompanies this plan.

7 Acknowledgments

Helpful comments on earlier drafts by John Van Sickle, Joan Baker, Selina Heppell, Paul Ringold, Bob Hughes, and Nathan Schumaker are gratefully acknowledged. Reviews by Thierry Oberdorff, Don McAllister, Yetta Jager, Dave Bradford, and Steve Ferraro were very helpful in improving the proposal.

8 References

Abell RA, Olson DM, Dinerstein E, Hurley PT, Diggs JT, Eichbaum W, Walters S, Wettengel W, Allnutt T, Loucks CJ, Hedao P. 1999. *Freshwater ecoregions of North America*. Island Press, Washington, DC. 368 pp.

Allen AP, Whittier TR, Larsen DP, Kaufmann PR, O'Connor RJ, Hughes RM, Stemberger RS, Dixit SS, Brinkhurst RO, Herlihy AT, Paulsen SG. (Submitted). Concordance of taxonomic composition patterns across multiple lake assemblages: effects of scale, body size, and land use. *Canadian Journal of Fisheries and Aquatic Sciences*.

Allen AP, Whittier TR, Kaufmann PR, Larsen DP, O'Connor RJ, Hughes RM, Stemberger RS, Dixit SS, Brinkhurst RO, Herlihy AT, Paulsen SG. 1999. Concordance of taxonomic richness patterns across multiple assemblages in lakes of northeastern USA. *Canadian Journal of Fisheries and Aquatic Sciences* 56: 739-747.

Angermeier PL, Schlosser IJ. 1989. Species-area relationships for stream fishes. *Ecology* 70(5): 1450-1462.

Angermeier PL, Schlosser IJ. 1995. Conserving aquatic biodiversity: beyond species and populations. *Evolution and the aquatic ecosystem: defining unique units in population conservation*. Neilsen JL, ed. American Fisheries Society Symposium 17, Bethesda, Maryland. pp. 402-416.

Angermeier PL, Winston MR. 1999. Characterizing fish community diversity across Virginia landscapes: prerequisite for conservation. *Ecological Applications* 9(1): 335-349.

Baker JP et al. (In preparation). *Scientific support for prioritizing wetland and stream/river corridor restoration to sustain aquatic and aquatic-dependent biota: a five-year research strategy.* US Environmental Protection Agency, Corvallis, Oregonpp.

Beecher HA, Dott ER, Fernau RF. 1988. Fish species richness and stream order in Washington State streams. *Environmental Biology of Fishes* 22(3): 193-209.

Billheimer D, Guttorp P, Fagan WF. (Submitted). Statistical analysis and interpretation of discrete compositional data. *Journal of the American Statistical Association*.

Bishop GD, Church MR, Daly C. 1998. Effects of improved precipitation estimates on automated runoff mapping: Eastern United States. *Journal of American Water Resources Association* 34(1): 159-166.

Böhning-Gaese K. 1997. Determinants of avian species richness at different spatial scales. *Journal of Biogeography* 24: 49-60.

Brown JH. 1981. Two decades of homage to Santa Rosalia: toward a general theory of diversity. *American Zoologist* 21: 877-888.

Brown JH. 1995. *Macroecology.* The University of Chicago Press, Chicago. 269 pp.

Brown JH, Lomolino MV. 1998. *Biogeography.* Sinauer Assoiciates, Inc., Sunderland, Massachusetts. 691 pp.

Chambers JM, Hastie TJ. 1992. *Statistical models in S*. Wadsworth & Brooks, Pacific Grove, CA. 608 pp.

Chorley RJ. 1969. The drainage basin as the fundamental geomorphic unit. *Introduction to physical hydrology.* Chorley RJ, ed. Methuen & Co, London. pp. 37-59.

Connell JH. 1978. Diversity in tropical rain forests and coral reefs. *Science* 199: 1302-1310.

Cook RE. 1969. Variation in species density of North American birds. *Systematic Zoology* 18: 63-84.

Cornell HV, Karlson RH. 1997. Local and regional precesses as controls of species richness. *Spatial ecology: the role of space in population dynamics and interspecific interactions*. Tilman D, Kareiva P, eds. Princeton University Press, Princeton. pp. 250-268.

Costa-Cabral MC, Burges SJ. 1997. Sensitivity of channel network planform laws and the question of topologic randomness. *Water Resources Research* 33(9): 2179-2197.

Currie DJ, Paquin V. 1987. Large-scale biogeographical patterns of species richness of trees. *Nature* 329: 326-327.

Currie DJ. 1991. Energy and large-scale patterns of animal- and plant-species richness. *American Naturalist* 137(1): 27-49.

Daly C, Neilson RP, Phillips DL. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meterology* 33: 140-158.

Davidowitz G, Rosenzweig ML. 1998. The latitudinal gradient of species diversity among North American grasshoppers (*Acrididae*) within a single habitat: a test of the spatial heterogeneity hypothesis. *Journal of Biogeography* 25: 553-560.

Drechsler M. 1998. Sensitivity analysis of complex models. *Biological Conservation* 86: 401-412.

Eadie JM, Hurly TA, Montgomerie RD, Teather KL. 1986. Lakes and rivers as islands: species-area relationships in the fish faunas of Ontario. *Environmental Biology of Fishes* 15(2): 81-89.

Edwards M, Morse DR. 1995. The potential for computer-aided identification in biodiversity research. *Trends in Ecology and Evolution* 10(4): 153-158.

Forman RTT, Godron M. 1986. *Landscape ecology.* John Wiley & Sons, New York. 619 pp.

Fortin M-J, Drapeau P. 1995. Delineation of ecological boundaries: comparison of approaches and significance tests. *Oikos* 72: 323-332.

Fortin M-J. 1999. Spatial statistics in landscape ecology. *Landscape ecological analysis: issues and applications*. Klopatek JM, Gardner RH, eds. Springer-Verlag, New York. pp. 253-279.

Foulds LR. 1992. *Graph theory applications*. Springer-Verlag, New York. 385 pp.

Fraser RH, Currie DJ. 1996. The species richness-energy hypothesis in a system where historical factors are thought to prevail: coral reefs. *American Naturalist* 148: 138-159.

Gaston KJ, Williams PH. 1993. Mapping the world's species--the higher taxon approach. *Biodiversity Letters* 1(1): 2-8.

Gerking SD. 1994. *Feeding ecology of fish*. Academic Press, San Diego. 416 pp.

Gore JA, Milner AM. 1990. Island biogeographical theory: can it be used to predict lotic recovery rates:. *Environmental Management* 14(5): 737-753.

Griffiths D. 1997. Local and regional species richness in North American lacustrine fish. *Journal of Animal Ecology* 66: 49-56.

Guégan J-F, Lek S, Oberdorff T. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* 391: 382-384.

Haggett P, Chorley RJ. 1969. *Network analysis in geography.* Edward Arnold, London. 348 pp.

Haining R. 1990. *Spatial data analysis in the social and environmental sciences*. Cambridge University Press, Cambridge. 409 pp.

Haydon D, Radtkey RR, Pianka ER. 1993. Experimental biogeography: interactions between stochastic, historical, and ecological processes in a model archipelago. *Species diversity in ecological communities: historal and geographical perspectives*. Ricklefs RE, Schluter D, eds. University of Chicago Press, Chicago. pp. 117-130.

Heck KL, Jr. 1976. Some critical considerations of the theory of species packing. *Evolutionary Theory* 1: 247-258.

Hengeveld R. 1990. *Dynamic biogeography.* Cambridge University Press, Cambridge. 249 pp.

Horsfield K, Woldenberg MJ. 1989. Diameters and cross-sectional areas of branches in the human pulmonary arterial tree. *The Anatomical Record* 223: 245-251.

Horwitz RJ. 1978. Temporal veriability patterns and the distributional patterns of stream fishes. *Ecological Monographs* 48: 307-321.

Hubbell SP. 1994. Towards a theory of biodiversity and biogeography on continuous landscapes. *Preparing for global change: a midwestern perspective*. Carmichael G, Folk G, Schnoor J, eds. SPB Academic Publishing, Amsterdam. pp. 171-199.

Hughes RM, Rexstad E, Bond CE. 1987. The relationship of aquatic ecoregions, river basins and physiographic provinces to the ichthyogeographic regions of Oregon. *Copeia* 1987: 423-432.

Hughes RM, Gammon JR. 1987. Longitudinal changes in fish assemblages and water quality in the Willamette River, Oregon. *Transactions of the American Fisheries Society* 116: 196-209.

Hughes RM, Kaufmann PR, Herlihy AT, Kincaid TM, Reynolds L, Larsen DP. 1998. A process for developing and evaluating indices of fish assemblage integrity. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 1618-1631.

Hugueny B. 1989. West African rivers as biogeographic islands: species richness of fish communities. *Oecologia* 79: 236-243.

Hugueny B, Paugy D. 1995. Unsaturated fish communities in African rivers. *American Naturalist* 146(1): 162-169.

Huston MA. 1994. *Biological diversity.* Cambridge University Press, Cambridge. 681 pp.

Hutchinson GE. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? *American Naturalist* 93: 145-159.

Jackson DA, Harvey HH. 1989. Biogeographic associations in fish assemblages: local vs. regional processes. *Ecology* 70(5): 1472-1484.

James FC, McCulloch CE. 1990. Multivariate analysis in ecology and systematics: panacea or Pandora's Box?? *Annual Review of Ecology and Systematics* 21: 129-166.

Jones KB, Riitters KH, Wickham JD, Tankersley RD, Jr., O'Neill RV, Chaloud DJ, Smith ER, Neale AC. 1997. An ecological assessment of the United States Mid-Atlantic region: a landscape atlas. *EPA/600/R-97/130*, US Environmental Protection Agency, Washington, DC.

Kaufman JH, Brodbeck D, Melroy OR. 1996. Critical biodiversity. *Conservation Biology* 12(3): 521-532.

Kelso JRM, Minns CK. 1996. Is fish species richness at sites in the Canadian Great Lakes the result of local or regional factors? *Canadian Journal of Fisheries and Aquatic Sciences* 53(Supplement 1): 175-193.

Kerr JT, Packer L. 1997. Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. *Nature* 385: 252-254.

Kiester AR. 1971. Species density of North American amphibians and reptiles. *Systematic Zoology* 20(2): 127-137.

Koopowitz H, Thornhill AD, Andersen M. 1994. A general stochastic model for the prediction of biodiversity losses based on habitat conversion. *Conservation Biology* 8(2): 425-438.

Legendre P, Legendre V. 1984. Postglacial dispersal of freshwater fishes in the Québec Peninsula. *Canadian Journal of Fisheries and Aquatic Sciences* 41: 1781-1802.

Lek S, Delacoste M, Baran P, Dimopoulos I, Lauga J, Aulagnier S. 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modeling* 90: 39-52.

Livingstone DA, Rowland M, Bailey PE. 1982. On the size of African riverine fish faunas. *American Zoologist* 22: 361-369.

Loveland TR, Merchant JW, Ohlen DO, Brown JF. 1991. Development of a land-dover characteristics database for the conterminous US. *Photogrammetric Engineering and Remote Sensing* 57: 1453-1463.

MacArthur RH, Wilson EO. 1967. *The theory of island biogeography.* Princeton University Press, Princetonpp.

Magnuson JJ, Benson BJ, McLain AS. 1994. Insights on species richness and turnover from long-term ecological research: fishes in north temperate lakes. *American Zoologist* 34: 437-451.

Maiorana VC. 1978. An explanation of ecological and developmental constants. *Nature* 273: 375-377.

Mandrak NE. 1995. Biogeographic patterns of fish species richness in Ontario lakes in relation to historical and environmental factors. *Canadian Journal of Fisheries and Aquatic Sciences* 52: 1462-1474.

Marks D. 1990. The sensitivity of potential evapogranspiration to climate change over the continental United States. *Biospheric feedbacks to climate change: the sensitivity of regional trace gas emissions, evapotranspiration, and energy balance to vegetation redistribution*. Gucinski H, Marks D, Turner DP, eds. US Environmental Protection Agency, Corvallis, OR. pp. IV.1-IV.31.

Master LL. 1991. Assessing threats and setting priorities for conservation. *Conservation Biology* 5(4): 559-563.

Master LL. 1994. Synoptic national assessment of comparative risks to biological diversity and landscape types: species distributions. , The Nature Conservancy, Arlington, Virginia.

Master LL. 1996. Predicting distributions of vertebrate species: some observations. *Gap analysis: a landscape approach to biodiversity planning*. Scott JM, Tear TH, Davis FW, eds. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland. pp. 171-176.

Master LL, Flack SR, Stein BA. 1998. Rivers of life: critical watersheds for protecting freshwater biodiversity. , The Nature Conservancy, Arlington, Virginia, 71.

Matthews RA, Matthews GB, Ehinger WJ. 1991. Classification and ordination of limnological data: a comparison of analytical tools. *Ecological Modelling* 53: 167-187.

Matthews WJ. 1998. *Patterns in freshwater fish ecology.* Chapman & Hall, New York. 756 pp.

Matuszek JE, Beggs GL. 1988. Fish species richness in relation to lake area, pH, and other abiotic factors in Ontario lakes. *Canadian Journal of Fisheries and Aquatic Sciences* 45: 1931-1941.

May RM, MacArthur RH. 1972. Niche overlap as a function of environmental variability. *Proceedings of the National Academy of Sciences* 69(5): 1100-1113.

McAllister DE, Platania SP, Schueler FW, Baldwin ME, Lee DS. 1986. Ichthyofaunal patterns on a geographic grid. *The zoogeography of North American freshwater fishes*. Hocutt CH, Wiley EO, eds. John Wiley & Sons, New York. pp. 17-51.

McAllister DE, Schueler FW, Roberts CM, Hawkins JP. 1994. Mapping and GIS analysis of the global distribution of coral reef fishes on an equal-area grid. *Mapping the diversity of nature*. Miller RL, ed. Chapman & Hall, London. pp. 155-175.

McCormick FH, Hughes RM, Kaufmann PR, Herlihy AT, Peck DV. (Submitted). Development of an index of biotic integrity for the Mid-Atlantic Highlands region. *Transactions of the American Fisheries Society.*

Meentemeyer V, Box EO. 1987. Scale effects in landscape studies. *Landscape heterogeneity and disturbance*. Turner MG, ed. Springer-Verlag, New York. pp. 15-34.

Moyle PB, Light T. 1996. Biological invasions of fresh water: empirical rules and assembly theory. *Biological Conservation* 78: 149-161.

Moyle PB, Light T. 1996. Fish invasions in California: do abiotic factors determine success? *Ecology* 77(6): 1666-1670.

Nielsen JL. 1995. Evolution and the aquatic ecosystem: defining unique units in population conservation. , American Fisheries Society Symposium 17, Bethesda, Maryland.

O'Connor RL, Jones MT, White D, Hunsaker C, Loveland T, Jones B, Preston E. 1996. Spatial partitioning of environmental correlates of avian biodiversity in the conterminous United States. *Biodiversity Letters* 3: 97-110.

O'Neill RV, DeAngelis DL, Waide JB, Allen TFH. 1986. *A hierarchical concept of ecosytems*. Princeton University Press, Princeton. 253 pp.

Oberdorff T, Guilbert E, Lucchetta J-C. 1993. Patterns of fish species richness in the Seine River basin, France. *Hydrobiologia* 259: 157-167.

Oberdorff T, Guégan J-F, Hugueny B. 1995. Global scale patterns of fish species richness in rivers. *Ecography* 18: 345-352.

Oberdorff T, Hugueny B, Guégan J-F. 1997. Is there an inflluence of historical events on contemporary fish species richness in rivers? Comparisons between Western Europe and North America. *Journal of Biogeography* 24: 461-467.

Oberdorff T, Hugueny B, Compin A, Belkessam D. 1998. Non-interactive fish communities in the coastal streams of North-western France. *Journal of Animal Ecology* 67: 472-484.

Omernik JM, Bailey RG. 1997. Distinguishing between watersheds and ecoregions. *Journal of the American Water Resources Association* 33(5): 935-949.

Osborne LL, Wiley MJ. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. *Canadian Journal of Fisheries and Aquatic Sciences* 49: 671-681.

Owen JG. 1989. Patterns of herpetofaunal species richness: relation to temperature, precipitation, and variance in elevation. *Journal of Biogeography* 16: 141-150.

Owen JG. 1990. Patterns of mammalian species richness in relation to temperature, productivity, and variance in elevation. *Journal of Mammalogy* 71(1): 1-13.

Owen JG. 1990. An analysis of the spatial structure of mammalian distribution patterns in Texas. *Ecology* 71(5): 1823-1832.

Phillips RB, Ehlinger TJ. 1995. Evolutionary and ecological considerations in the reestablishment of Great Lakes Coregonid fishes. *Evolution and the aquatic ecosystem: defining unique units in population conservation*. Neilsen JL, ed. American Fisheries Society Symposium 17, Bethesda, Maryland. pp. 133-144.

Pianka ER. 1978. *Evolutionary ecology.* Harper & Row, New York. 397 pp.

Pickett ST, Kolasa J, Jones CG. 1994. *Ecological Understanding*. Academic Press, San Diego. 206 pp.

Rahbek C. 1995. The elevational gradient of species richness: a uniform pattern? *Ecography* 18(2): 200-205.

Rappoldt C, Hogeweg P. 1980. Niche packing and number of species. *American Naturalist* 116(4): 480-492.

Rathert D, White D, Sifneos JC, Hughes RM. 1999. Environmental correlates of species richness for native freshwater fish in Oregon, USA. *Journal of Biogeography* 26(2): 257-274.

Reed BC, Brown JF, VanderZee D, Loveland TR, Merchant JW, Ohlen DO. 1994. Measuring phenological variability from satellite. *Journal of Vegetation Science* 5: 703-714.

Reiss MJ. 1989. *The allometry of growth and reproduction*. Cambridge University Press, Cambridge. 182 pp.
Reynolds JH, Ford ED. 1999. Multi-criteria assessment of ecological process models. *Ecology* 80(2): 538-553.
Ricklefs RE. 1987. Community diversity: relative roles of local and regional processes. *Science* 235: 167-171.
Ricklefs RE, Schluter D, eds. 1993. *Species diversity in ecological communities: historical and geographical perspectives*. University of Chicago Press, Chicago. 414 pp.
Rogers JS. 1976. Species density and taxonomic diversity of Texas amphibians and reptiles. *Systematic Zoology* 25: 26-40.
Rosenzweig ML, Abramsky Z. 1993. How are diversity and productivity related? *Species diversity in ecological communities: historal and geographical perspectives*. Ricklefs RE, Schluter D, eds. University of Chicago Press, Chicago. pp. 52-65.
Rosenzweig ML. 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge. 436 pp.
Roth VL. 1981. Constancy in the size ratios of sympatric species. *American Naturalist* 118: 394-404.
Schumaker NH. 1998. A user's guide to the PATCH model. *EPA/600/R-98/135*, US Environmental Protection Agency, Corvallis, Oregon.
Seaber PR, Kapinos FP, Knapp GL. 1987. Hydrologic unit maps. *US Geological Survey Water-Supply Paper 2294*, US Government Printing Office, Washington.
Seagle SW, Shugart HH. 1985. Faunal richness and turnover on dynamic landscapes: a simulation study. *Journal of Biogeography* 12: 499-508.
Seagle SW. 1986. Generation of species-area curves by a model of animal-habitat dynamics. *Wildlife 2000: modeling habitat relationships of terrestrial vertebrates*. Verner J, Morrison ML, Ralph CJ, eds. University of Wisconsin Press, Madison. pp. 281-285.
Shrader-Frechette KS, McCoy ED. 1993. *Method in ecology.* Cambridge University Press, Cambridge. 328 pp.
Shvarts EA, Pushkaryov SV, Krever VG, Ostrovsky MA. 1995. Geography of mammal diversity and searching for ways to predict global changes in biodiversity. *Journal of Biogeography* 22: 907-914.
Sifneos JC, White D, Urquhart NS. (Submitted). A comparison of pruning methods for regression trees: evidence from simulation and published studies. *Journal of Agricultural, Biological, and Environmental Statistics*.
Simberloff D, Boecklen W. 1981. Santa Rosalia reconsidered: size ratios and competition. *Evolution* 35: 1206-1228.
Simpson GG. 1964. Species density of North American recent mammals. *Systematic Zoology* 13: 57-73.
Smith GR. 1981. Late Cenozoic freshwater fishes of North America. *Annual Review of Ecology and Systematics* 12: 163-193.
Stein BA, Flack SR. 1997. 1997 species report card: the state of US plants and animals. , The Nature Conservancy, Arlington, Virginia.
USEPA. 1990. Reducing risks: setting priorities and strategies for environmental protection. *SAB-EC-90-021*, Science Advisory Board. US Environmental Protection Agency, Washington, DC.
USSR. 1977. *Atlas of world water balance*. USSR National Committee for the International Hydrological Decade. United Nations Educational, Scientific, and Cultural Organization, Paris, Francepp.
Venables WN, Ripley BD. 1997. *Modern applied statistics with S-Plus, 2nd edition*. Springer-Verlag, New York. 568 pp.
Villa F, Rossi O, Sartore F. 1992. Understanding the role of chronic environmental disturbance in the context of island biogeographic theory. *Environmental Management* 16(5): 653-666.
Waples RS. 1995. Evolutionarily significant units and the conservation of biological diversity under the Endangered Species Act. *Evolution and the aquatic ecosystem: defining unique units in population conservation*. Neilsen JL, ed. American Fisheries Society Symposium 17, Bethesda, Maryland. pp. 8-27.
Ward JV. 1998. Riverine landscapes: biodiversity patterns, disturbance regimes, and aquatic conservation. *Biological Conservation* 83(3): 269-278.
White D, Kimerling AJ, Overton WS. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. *Cartography and Geographic Information Systems* 19(1): 5-22.
White D, Freemark KE, Preston EM, Kiester AR. 1999. A hierarchical framework for conserving biodiversity. *Landscape ecological analysis: issues and applications*. Klopatek JM, Gardner RH, eds. Springer-Verlag, New York. pp. 127-153.
Whittier TR, Kincaid TM. (In press). Introduced fish in northeast USA lakes: regional extent, dominance, and effects on native species richness. *Transactions of the American Fisheries Society.*
Whittier TR, Hartel KE. 1997. First records of Redear Sunfish (*Lepomis microlophus*) in New England. *Northeastern Naturalist* 4(4): 237-240.

Whittier TR, Halliwell DB, Paulsen SG. 1997. Cyprinid distributions in northeast USA lakes: evidence of regional scale biodiversity losses. *Canadian Journal of Fisheries and Aquatic Sciences* 54: 1593-1607.

Wickham JD, Wu J, Bradford DF. 1997. A conceptual framework for selecting and analyzing stressor data to study species richness at large spatial scales. *Environmental Management* 21: 247-257.

Williams PH, Gaston KJ. 1994. Measuring more of biodiversity: can higher-taxon richness predict wholesale species richness? *Biological Conservation* 67: 211-217.

Wilson JW, III. 1974. Analytical zoogeography of North American mammals. *Evolution* 28: 124-140.

Winemiller KO. 1995. Aspects structurels et fonctionnels de la biodiversité des peuplements de poissons. *Bulletin Français de la Pêche et de la Pisciculture* 337/338/339: 23-45.

Withers MA, Meentemeyer V. 1999. Concepts of scale in landscape ecology. *Landscape ecological analysis: issues and applications*. Klopatek JM, Gardner RH, eds. Springer-Verlag, New York. pp. 205-252.

Wootton RJ. 1990. *Ecology of teleost fishes*. Chapman & Hall, London. 404 pp.

Wright DH. 1983. Species-energy theory: an extension of species-area theory. *Oikos* 41: 496-506.

Wright DH, Currie DJ, Maurer BA. 1993. Energy supply and patterns of species richness on local and regional scales. *Species diversity in ecological communities: historal and geographical perspectives*. Ricklefs RE, Schluter D, eds. University of Chicago Press, Chicago. pp. 66-74.

Ziv Y. 1998. The effect of habitat heterogeneity on species diversity patterns: a community-level approach using an object-oriented landscape simulation model (SHALOM). *Ecological Modeling* 111: 135-170.